AN INTRODUCTION TO
STATISTICAL
THERMODYNAMICS

by

TERRELL L. HILL

Department of Chemistry
University of Oregon

x

properties of a system containing many molecules, as a starting point we must certainly be provided with knowledge of the properties of the individual molecules making up the system and of the nature of the interactions between these molecules. This is information which can in principle be furnished by quantum mechanics but which in practice is usually obtained from experiments based on the behavior of individual molecules (e.g., spectroscopy), pairs of molecules (e.g., the second virial coefficient of an imperfect gas), etc.

Although quantum mechanics is prerequisite to statistical mechanics, fortunately a reasonably satisfactory version of statistical mechanics can be presented without using any quantum-mechanical concepts other than those of quantum-mechanical states, energy levels, and intermolecular forces. Only in Part IV of the book is it necessary to go beyond this stage.

Another very helpful simplification is that the classical limit of quantum mechanics can be used, without appreciable error, in most problems involving significant intermolecular interactions. Problems of this type are very difficult without this simplification (Part IV).

Despite our extensive use of classical statistical mechanics in the applications of Parts II and III, we introduce the principles of statistical mechanics, beginning in the next section, in quantum-mechanical language because the argument is not only more general but is actually much simpler this way.

1–2 Ensembles and postulates. As mentioned above, our problem is to calculate macroscopic properties from molecular properties. Our general approach is to set up postulates which allow us to proceed directly with this task insofar as "mechanical" thermodynamic properties are concerned; the "nonmechanical" properties are then handled indirectly by an appeal to thermodynamics. By "mechanical" properties we mean, for example, pressure, energy, volume, number of molecules, etc., all of which can be defined in purely mechanical terms (quantum or classical) without, for example, introducing the concept of temperature. Examples of "nonmechanical" thermodynamic variables are temperature, entropy, free energy (Gibbs or Helmholtz), chemical potential, etc.

Let us consider the pressure as a typical mechanical variable. In principle, if we wished to calculate the pressure in a thermodynamic system from molecular considerations, we would have to calculate (by quantum or possibly classical mechanics) the force per unit area exerted on a wall of the system, taking into account the change in the state of the whole system with time. The force itself would be a function of time. What we would need, therefore, is a time average of the force over a period of time sufficiently long to smooth out fluctuations, i.e., sufficiently long to give

a time average which is independent, say, of the starting time, $t = t_0$, in the averaging. Because of the tremendous number of molecules in a typical system, and the fact that they interact with each other, such a hypothetical calculation is of course completely out of the question in either quantum or classical mechanics.

Therefore we are forced to turn to an alternative procedure, the ensemble method of Gibbs, based on postulates connecting the desired time average of a mechanical variable with the ensemble average (defined below) of the same variable. The validity of these postulates rests on the agreement between experiment and deductions (such as those in this book) made from the postulates. So far, there is no experimental evidence available that casts doubt on the correctness of the postulates of statistical mechanics.

Before stating the postulates, we must introduce the concept of an ensemble of systems. An ensemble is simply a (mental) collection of a very large number $\mathfrak{N}$ of systems, each constructed to be a replica on a thermodynamic (macroscopic) level of the actual thermodynamic system whose properties we are investigating. For example, suppose the system of interest has a volume V, contains N molecules of a single component, and is immersed in a large heat bath at temperature T. The assigned values of N, V, and T are sufficient to determine the thermodynamic state of the system. In this case, the ensemble would consist of $\mathfrak{N}$ systems, all of which are constructed to duplicate the thermodynamic state (N, V, T) and environment (closed system immersed in a heat bath) of the original system. Although all systems in the ensemble are identical from a thermodynamic point of view, they are not all identical on the molecular level. In fact, in general, there is an *extremely* large number of quantum (or classical) states consistent with a given thermodynamic state. This is to be expected, of course, since three numbers, say the values of N, V, and T, are quite inadequate to specify the detailed molecular (or "microscopic") state of a system containing something in the order of 10^{20} molecules.

Incidentally, when the term "quantum state" is used here, it will be understood that we refer specifically to *energy* states (i.e., energy eigenstates, or stationary states).

At any instant of time, in an ensemble constructed by replication of a given thermodynamic system in a given environment, many different quantum states are represented in the various systems of the ensemble. In the example mentioned above, the calculated instantaneous pressure would in general be different in these different quantum states. The "ensemble average" of the pressure is then the average over these instantaneous values of the pressure, giving the same weight to *each system* in the ensemble in calculating the average. A similar ensemble average can be calculated for any mechanical variable which may have different values

(i.e., which is not held constant) in the different systems of the ensemble.

We now state our *first postulate: the (long) time average of a mechanical variable M in the thermodynamic system of interest is equal to the ensemble average of M, in the limit as $\mathfrak{N} \rightarrow \infty$, provided that the systems of the ensemble replicate the thermodynamic state and environment of the actual system of interest.* That is, this postulate tells us that we may replace a time average on the one actual system by an instantaneous average over a large number of systems "representative" of the actual system. The first postulate by itself is not really helpful; we need in addition, in order to actually compute an ensemble average, some information about the relative probability of occurrence of different quantum states in the systems of the ensemble. This information must be provided in a second postulate.

Note that the ensemble average of M in the limit as $\mathfrak{N} \rightarrow \infty$, referred to above, must be independent of time. Otherwise the original system which the ensemble "represents" is not in equilibrium.

We shall work out details in this chapter for the three most important thermodynamic environments: (a) *an isolated system* (N, V, and E given, where $E =$ energy); (b) *a closed, isothermal system* (N, V, and T given); and (c) *an open, isothermal system* (μ, V, and T given, where $\mu =$ chemical potential). N and μ stand for the sets N_1, N_2, ... and μ_1, μ_2, ... if the system contains more than one component. Also, V might stand for a set of "external variables"* if there are more than one. The representative ensembles in the above three cases are usually called *microcanonical*, *canonical*, and *grand canonical*, respectively. The first postulate is applicable to all these cases and to other ensembles which will be introduced in Section 1–7. The second postulate, however, can be limited to a statement concerning only the microcanonical ensemble. The corresponding statement for other ensembles can then be deduced (as in Section 1–3, for example) from this limited second postulate without any further assumptions.

Our *second postulate* is: *in an ensemble* ($\mathfrak{N} \rightarrow \infty$) *representative of an isolated thermodynamic system, the systems of the ensemble are distributed uniformly, that is, with equal probability or frequency, over the possible quantum states consistent with the specified values of N, V, and E.* In other words, each quantum state is represented by the same number of systems in the ensemble; or, if a system is selected at random from the ensemble, the probability that it will be found in a particular quantum state is the same for all the possible quantum states. A related implication of this postulate, when combined with the first postulate, is that the single isolated system of actual interest (which serves as the prototype for the

---

* There is one "external variable" for each thermodynamic work term, e.g., volume, area, length, etc.

systems of the ensemble) spends equal amounts of time, over a long period of time, in each of the available quantum states. This last statement is often referred to as the quantum "ergodic hypothesis," while the second postulate by itself is usually called the "principle of equal *a priori* probabilities." The ergodic hypothesis in classical statistical mechanics is mentioned at the end of Section 6–3. For a more detailed discussion, see Tolman, pp. 63–70 and 356–361. (For full identification of works referred to by only the author's last name, see Preface.)

The value of E in the second postulate must be one of the energy levels of the quantum-mechanical system defined by N and V. Since N is extremely large, the energy levels for such a system will be so close together as to be practically continuous, and furthermore, each of these levels will have an extremely high degeneracy. We shall in general denote the number of quantum states (i.e., the degeneracy) associated with the energy level E for a quantum-mechanical system with N and V by $\Omega(N, V, E)$. Thus the number of "possible quantum states" referred to in the second postulate is Ω.

A complication in the above discussion is the fact that, from an operational point of view, E cannot be known precisely; there will always be a small uncertainty δE in the value of E. For all thermodynamic purposes this complication is completely inconsequential.* Hence for the sake of simplicity we ignore it.

It should also be mentioned that the point of view in the above statement of the second postulate is not so general as it might be. If the energy level E for the system N, V has a degeneracy Ω, there are Ω orthogonal (and therefore linearly independent) wave functions ψ which satisfy the Schrödinger equation $\mathcal{H}\psi = E\psi$. The particular choice of the Ω ψ's is somewhat arbitrary, since other possible choices can always be set up by forming suitable linear combinations of the ψ's in the first choice. In any case, the "Ω quantum states" mentioned in connection with the second postulate refers to some set of orthogonal ψ's all "belonging" to the same E. But regardless of the set of ψ's chosen, the wave function representing the actual quantum-mechanical state of any system selected from the ensemble will in general *not* be one of the chosen set of ψ's, but will be some linear combination of all of them. The contrary is really implied in the above statement of the second postulate. Fortunately, this simplification in our statement of the postulate makes no difference† in any deductions we shall make that can be compared with experiment. In fact, this complication can be bypassed, at the expense of slightly more complicated

---

* See S. M. (the present author's earlier work identified in the Preface), p. 113, and Mayer and Mayer, pp. 55–56, 100–102.

† See S. M., pp. 50–55, 79.

equations in the next section, by using energy *levels* (rather than *states*) in the postulate, as in Eq. (1–38) (Problem 1–1).

We turn now to a derivation from the above two postulates of the essential properties of the canonical and grand ensembles.

1–3 Canonical ensemble. The experimental system of interest here has a fixed volume V, fixed numbers of molecules N (which stands for N_1, N_2, ... in a multicomponent system), and is immersed in a very large heat bath at temperature T. The heat bath is assumed "very large" to be consistent with the use of the limit $\mathfrak{N} \to \infty$ below. Our first objective is to set up the machinery necessary for calculating the average value of mechanical variables, such as energy and pressure, in the system. In view of the first postulate, this means that we need to be able to calculate the *ensemble* average of such variables. This, in turn, can be done if we know the value of the particular variable in question in a given quantum state and the fraction of systems in the ensemble which are in this quantum state. It might be noted that because the thermodynamic system here is not isolated but is in contact with a heat bath, the energy of the system can fluctuate; therefore quantum states belonging to *different* energy levels E will have to be reckoned with. Since mechanical variables have well-defined values in a given quantum state (in fact we can use this property as the definition of a "mechanical variable"), the task that remains is to determine the fraction of systems in the ensemble in a given quantum state (or the probability that a system selected arbitrarily from the ensemble is in a given quantum state). This is the problem we now consider.

The experimental, or prototype, system is in a very large heat bath at temperature T. Therefore each system in the ensemble representative of the experimental system must also be in a very large heat bath at T. Specifically, we contemplate the following arrangement, which satisfies this requirement. We imagine $\mathfrak{N}$ macroscopic systems as our ensemble, each with N and V (duplicating the values in the experimental system), stacked together in a lattice (Fig. 1–1). The walls between the different systems in the ensemble are heat conducting, but impermeable to all molecules. To establish the temperature T, we imagine further that the entire stack of systems (i.e., the ensemble) is placed in a sufficiently large heat bath at T. After equilibrium is reached, thermal insulation (represented schematically by the double lines in Fig. 1–1) is placed on the outside walls of the ensemble, and the ensemble is removed from the heat bath. The entire ensemble itself is now an isolated system with volume $\mathfrak{N}V$, numbers of molecules $\mathfrak{N}N$, and a total energy which we shall denote by E_t (t = total). The relation between E_t and the temperature T will emerge later. Observe that each system in the ensemble is immersed in a large (we shall later use the limit $\mathfrak{N} \to \infty$) heat bath at temperature T,
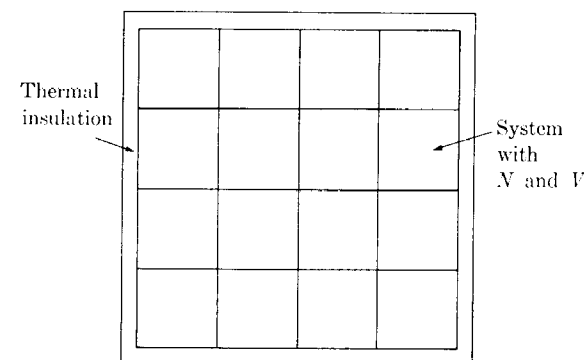
FIG. 1–1. Canonical ensemble of $\mathfrak{N}$ systems, each with N and V.

as is required if the ensemble is to be representative of the original thermodynamic system. That is, the remaining $\mathfrak{N} - 1$ systems in the ensemble serve as the heat bath for any one selected system.

At this point we come to the essential step in the argument, which is to note that since the ensemble itself is an isolated system, *we can apply the second postulate to the whole ensemble.* Thus the entire canonical ensemble shown in Fig. 1–1 is now regarded as a prototype thermodynamic system, characterized by the variables $\mathfrak{N}V$, $\mathfrak{N}N$, and E_t. We shall refer to this system as a "supersystem" in order to avoid confusion with the original experimental closed, isothermal system. The second postulate tells us, then, that every possible quantum state of this supersystem (canonical ensemble) is equally probable and hence should be given equal weight in the calculation of average values of interest. As we show next, it is possible to take advantage of this observation in order to find the required probability of occurrence of a given quantum state in the systems of a canonical ensemble.

We return now to a single system in the canonical ensemble. As a quantum-mechanical system, it is characterized by N and V. Let us list all possible energy states for such a system in increasing order of the energy eigenvalue, $E_1, E_2, \ldots, E_j, \ldots$. Here, for later convenience, each state is listed separately so that when degeneracy occurs several successive E_j's will have the same value. For example, in the notation used above and to which we shall return later, the energy value E occurs Ω successive times in the list.

Each E_j is a function of N and V. If V is changed infinitely slowly, each E_j changes in a continuous manner. However, the number of molecules of any one of the components can be changed only discontinuously—one molecule at a time. Hence the energy levels must jump discontinuously if N changes.

For most systems containing many molecules, it is not possible, for purely mathematical reasons, to actually calculate the energies E_1, E_2, ... from Schrödinger's equation. But for generality, we assume in the present argument that these energies are known. In applications, we shall be able to make progress in each case only to the extent that we can overcome this difficulty either by approximation, or by use of classical mechanics, or by reducing this many-body problem to a one-, two-, three-, ... body problem, etc. In any case, the ensemble method has the following advantage over a direct time-average calculation on a single system (see Section 1–2): we need only the stationary states of the system and do not have to follow the change in state of the system with time.

The list of energy eigenvalues E_1, E_2, ... is assumed, then, to be the correct list for any given problem. The argument that follows is valid irrespective of such complications as intermolecular forces, symmetry restrictions on wave functions, etc.

Since each system in the canonical ensemble has the same N and V, all systems have the same set of energy states, represented by E_1, E_2, ..., E_j, ... Now suppose we observe, simultaneously, the energy state of each system in the ensemble, and count the number of systems found in each of the listed states. We let n_1 be the number of systems found in state E_1, ..., n_j in state E_j, etc. The set of numbers n_1, n_2, ... is called a "distribution." There are, of course, many possible distributions that might be observed, but obviously all must satisfy the relations

$$\sum_j n_j = \mathfrak{N}, \tag{1-1}$$

$$\sum_j n_j E_j = E_t. \tag{1-2}$$

The individual systems in the supersystem (canonical ensemble) are macroscopic in size, are arranged in a certain order, and can be separately labeled. Then the energy state of the whole supersystem would be completely specified if we indicated the system energy state (i.e., E_1, E_2, ...) for each of the (labeled) systems in the supersystem. To take a simple example, suppose there are four systems (A, B, C, D) in the supersystem ($\mathfrak{N} = 4$) and the possible energy states for each system are E_1, E_2, and E_3. Then one possible energy state for the supersystem would be, say,

A	B	C	D
E_2	E_3	E_2	E_1

provided that (compare Eq. 1–2)

$$E_1 + 2E_2 + E_3 = E_t \text{ (preassigned)}.$$

Here $n_1 = 1$, $n_2 = 2$, $n_3 = 1$. Actually, there are 12 possible states of the supersystem consistent with this distribution. Three of them are:

A	B	C	D
E_2	E_2	E_3	E_1
E_2	E_3	E_2	E_1
E_3	E_2	E_2	E_1

But there are four sets of this type, corresponding to the four possible assignments of E_1. In general, the number of states of the supersystem, $\Omega_t(n)$, consistent with a given distribution n_1, n_2, ... (n represents the entire set n_1, n_2, ...) is given by the well-known combinatorial formula

$$\Omega_t(n) = \frac{(n_1 + n_2 + \cdots)!}{n_1! n_2! \cdots} = \frac{\mathfrak{N}!}{\prod_j n_j!}. \tag{1-3}$$

Recall that we are attempting to find the probability of observing a given quantum state (say E_j) in a system selected from a canonical ensemble (or the fraction of systems in the ensemble in the state E_j). For a particular distribution n_1, n_2, ..., this probability or fraction is just $n_j/\mathfrak{N}$ for state E_j. But, in general, there are very many possible distributions for given N, V, $\mathfrak{N}$, and E_t. What we need is the over-all probability; that is, an average of $n_j/\mathfrak{N}$ over these distributions, based on an assignment of equal weight to each state of the supersystem. Assignment of equal weights to supersystem states implies immediately that the weight assigned to each distribution, in calculating an average over different distributions, should be proportional to $\Omega_t(n)$ for the distribution.

Now consider the numerical example above, and suppose further that there are just two distributions which satisfy the conditions of Eqs. (1–1) and (1–2), namely,

$$n_1 = 1, n_2 = 2, n_3 = 1, \quad \Omega_t = 12,$$
$$n_1 = 2, n_2 = 0, n_3 = 2, \quad \Omega_t = 6.$$

The probability of observing E_3 is $\frac{1}{4}$ in the first distribution and $\frac{1}{2}$ in the second distribution, while the over-all probability is $\frac{1}{3}$:

$$\bar{n}_3 = \frac{1 \times 12 + 2 \times 6}{12 + 6} = \frac{4}{3}, \quad \frac{\bar{n}_3}{\mathfrak{N}} = \frac{1}{3}.$$

In general, the required probability of observing a given quantum state E_j in an arbitrary system of a canonical ensemble is

$$P_j = \frac{\bar{n}_j}{\mathfrak{N}} = \frac{1}{\mathfrak{N}} \frac{\sum_n \Omega_t(n) n_j(n)}{\sum_n \Omega_t(n)}, \tag{1-4}$$

where $n_j(n)$ means the value of n_j in the distribution n. The sum is over all distributions satisfying Eqs. (1–1) and (1–2). Of course, by definition, $\sum_j P_j = 1$.

Then the desired ensemble averages of, for example, the energy and pressure are

$$\overline{E} = \sum_j P_j E_j \qquad (1\text{–}5)$$

and

$$\overline{p} = \sum_j P_j p_j, \qquad (1\text{–}6)$$

where p_j is the pressure in state E_j, defined by

$$p_j = -\left(\frac{\partial E_j}{\partial V}\right)_N. \qquad (1\text{–}7)$$

That is, $-p_j\,dV = dE_j$ is the work that has to be done on the system, when in the state E_j, in order to increase the volume by dV.

In principle, Eq. (1–4) for P_j tells us all we need to know to calculate canonical ensemble averages of mechanical variables. But in practice, a much more explicit expression for P_j is necessary. We must now face this problem.

The most elegant way to proceed is to employ the Darwin-Fowler technique,* based on the mathematical method of steepest descents. However, in the present discussion, since we can take $\mathfrak{N} \to \infty$, the so-called maximum-term method, which involves the use of undetermined multipliers, is equally rigorous though not so elegant. The latter method, which we shall use, has the important advantage of requiring much less of the reader in the way of mathematical background.

In any particular case we are given $\mathfrak{N}$, the E_j (determined by N and V), and E_t (determined by $\mathfrak{N}$, N, V, and T). There are then many possible distributions n consistent with the restrictions of Eqs. (1–1) and (1–2). For each of these distributions we can calculate from Eq. (1–3) the weight $\Omega_t(n)$ to be used in obtaining averages, as already explained. The situation here parallels exactly that illustrated in Appendix II. That is, because of the large numbers involved (the present example is ideal in this respect because we can take the limit $\mathfrak{N} \to \infty$), the most probable distribution, and distributions which differ only negligibly from the most probable distribution, completely dominate the computation of the average in Eq. (1–4). By the most probable distribution, denoted by n^*, we mean of course that distribution to which the largest $\Omega_t(n)$ belongs. In effect this means that, in the limit as $\mathfrak{N} \to \infty$, we can regard all other weights $\Omega_t(n)$
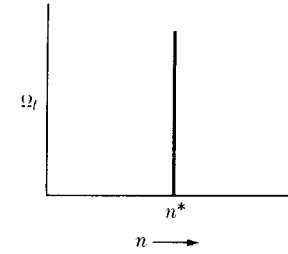
---

* See, for example, Schrödinger, Chapter 6.

FIG. 1–2. Number of states Ω_t as a function of the distribution n (schematic).

as negligible compared with $\Omega_t(n^*)$. This is illustrated diagrammatically in Fig. 1–2. With $\mathfrak{N}$ large but finite, there would be a narrow gaussian distribution centered about $n = n^*$. But in the limit as $\mathfrak{N} \to \infty$, this distribution becomes completely sharp (a Dirac δ-function).

Naturally, as we let $\mathfrak{N} \to \infty$ (i.e., increase the size of the ensemble), holding N, V, and T fixed, each $n_j \to \infty$ also. But all ensemble averages depend only on the ratio $n_j/\mathfrak{N}$, which remains finite.

Equation (1–4) becomes, then,

$$P_j = \frac{\overline{n}_j}{\mathfrak{N}} = \frac{1}{\mathfrak{N}} \frac{\Omega_t(n^*)n_j^*}{\Omega_t(n^*)} = \frac{n_j^*}{\mathfrak{N}}, \qquad (1\text{–}8)$$

where n_j^* is the value of n_j in the most probable distribution, n^*. Equation (1–8) tells us that in the computation of P_j we can replace the mean value of n_j by the value of n_j in the most probable (largest Ω_t) distribution. This leads us to a purely mathematical question: Which of all possible sets of n_j's satisfying Eqs. (1–1) and (1–2) gives us the largest Ω_t?

We solve this problem by the method of undetermined multipliers (see Appendix III). The distribution giving the largest Ω_t is also the distribution giving the largest $\ln \Omega_t$, since $\ln x$ increases monotonically with x. We work with $\ln \Omega_t$ instead of Ω_t because it is more convenient. From Eq. (1–3),

$$\ln \Omega_t(n) = \left(\sum_i n_i\right) \ln \left(\sum_i n_i\right) - \sum_i n_i \ln n_i,$$

where we have used Stirling's approximation (Appendix II) and changed the running index from j to i. This "approximation" is in fact exact here because we are interested in the limit $\mathfrak{N}, n_i \to \infty$. According to the method of undetermined multipliers, the set of n_j's which leads to the maximum value of $\ln \Omega_t(n)$, subject to the conditions (1–1) and (1–2), is found from the equations

$$\frac{\partial}{\partial n_j}\left[\ln \Omega_t(n) - \alpha \sum_i n_i - \beta \sum_i n_i E_i\right] = 0, \qquad j = 1, 2, \ldots,$$

where α and β are the undetermined multipliers. On carrying out the differentiation, we find

$$\ln\left(\sum_i n_i\right) - \ln n_j^* - \alpha - \beta E_j = 0, \qquad j = 1, 2, \ldots,$$

or

$$n_j^* = \mathfrak{N}e^{-\alpha}e^{-\beta E_j}, \qquad j = 1, 2, \ldots. \qquad (1\text{--}9)$$

This is the most probable distribution, expressed in terms of α and β. If desired, $\mathfrak{N}$ may be substituted for $\sum_i n_i$ in $\ln \Omega_t(n)$ at the outset, and treated as a constant in the differentiation. This will change the meaning of α, but not any physical results.

The straightforward procedure here is to substitute the distribution (1–9) into Eqs. (1–1) and (1–2) in order to determine α and β as functions of $\mathfrak{N}$ and E_t, or of $\mathfrak{N}$ and $\overline{E}$ (since obviously $E_t = \mathfrak{N}\overline{E}$). The result is

$$e^\alpha = \sum_j e^{-\beta E_j}, \qquad (1\text{--}10)$$

$$\overline{E} = \frac{\sum_j E_j e^{-\beta E_j}}{\sum_j e^{-\beta E_j}}, \qquad (1\text{--}11)$$

where $\mathfrak{N}$ has dropped out of both equations. Equation (1–11) provides β as an implicit function of $\overline{E}$ (and also of N and V, since the energies E_j are functions of N and V). Equation (1–10) then gives α in terms of β (and N, V). However, the independent variables of real interest here are N, V, T rather than $N, V, \overline{E}$, and we have no information yet about the dependence of $\overline{E}$ on T. Hence we do not pursue the above approach any further (see Problem 1–2, however), but turn instead, in Section 1–4, to a thermodynamic argument which provides a direct connection between β and T.

We note in passing that elimination of $e^{-\alpha}$ in Eq. (1–9) by use of Eq. (1–10) (or comparison of Eqs. 1–5 and 1–11) gives us P_j as a function of β, N, and V:

$$P_j = \frac{n_j^*}{\mathfrak{N}} = \frac{e^{-\beta E_j(N,V)}}{\sum_i e^{-\beta E_i(N,V)}}, \qquad j = 1, 2, \ldots. \qquad (1\text{--}12)$$

Anticipating the fact that β turns out to be a positive number, we deduce from this equation that the probability of observing a given quantum state in a canonical ensemble decreases exponentially with the energy of the quantum state.

1–4 Canonical ensemble and thermodynamics. To bring nonmechanical thermodynamic variables such as temperature and entropy into our dis-

cussion, we now combine the above "mechanical" considerations with thermodynamics. In the first place, by virtue of the first postulate, we can associate the thermodynamic pressure p and energy E with the statistical-mechanical ensemble averages $\overline{p}$ and $\overline{E}$. Let us take the differential of $\overline{E}$ in Eq. (1–5), holding N constant (the system being closed):

$$d\overline{E} = \sum_j E_j \, dP_j + \sum_j P_j \, dE_j$$

$$= -\frac{1}{\beta}\sum_j (\ln P_j + \ln Q)\, dP_j + \sum_j P_j \left(\frac{\partial E_j}{\partial V}\right)_N dV, \quad (1\text{--}13)$$

where we have defined

$$Q = \sum_j e^{-\beta E_j}, \qquad (1\text{--}14)$$

used Eq. (1–12) in the first sum, and have recognized in the second sum that $E_j(N, V)$ can vary only with V if N is fixed. The first sum simplifies further in view of the relations

$$\sum_j P_j = 1, \qquad \sum_j dP_j = 0,$$

and

$$d\left(\sum_j P_j \ln P_j\right) = \sum_j \ln P_j \, dP_j.$$

Thus, using Eq. (1–6), we can write

$$-\frac{1}{\beta} d\left(\sum_j P_j \ln P_j\right) = d\overline{E} + \overline{p}\, dV. \qquad (1\text{--}15)$$

Since we already have the associations with thermodynamics $E \leftrightarrow \overline{E}$ and $p \leftrightarrow \overline{p}$, and since in thermodynamics (N constant)

$$T\, dS = dE + p\, dV,$$

we can deduce from Eq. (1–15) the further association

$$T\, dS \leftrightarrow -\frac{1}{\beta} d\left(\sum_j P_j \ln P_j\right). \qquad (1\text{--}16)$$

With these associations established, let us digress to note that from Eq. (1–13) and

$$dE = DQ^* - DW,$$

we have

$$DQ^* = T\,dS \leftrightarrow \sum_j E_j\,dP_j, \qquad (1\text{--}17)$$

$$DW = p\,dV \leftrightarrow \sum_j P_j\,dE_j, \qquad (1\text{--}18)$$

where Q^* and W are heat absorbed and work done by the system, respectively. These relations provide us, in a general way, with the molecular interpretation of the thermodynamic concepts of heat and work. We see that when a closed thermodynamic system increases its energy infinitesimally by the absorption of heat from its surroundings, this is accomplished not by changing the energy levels of the system but rather by a shift in the fraction of time the system spends in the various energy states. The converse statement can be made about the work term.

We now return to the main argument, the purpose of which is to relate S to the P_j. From Eq. (1–16),

$$dS \leftrightarrow \frac{1}{\beta T}\,dG, \qquad (1\text{--}19)$$

where G is defined by

$$G = -\sum_j P_j \ln P_j.$$

From thermodynamics we know that the left side of Eq. (1–19) is an exact differential. Hence the right side must be also. This condition will be met provided that $1/\beta T$ is any function of G, say $\varphi(G)$. That is,

$$dS \leftrightarrow \varphi(G)\,dG = df(G), \qquad (1\text{--}20)$$

where

$$f(G) = \int \varphi(G)\,dG, \qquad \varphi(G) = \frac{df(G)}{dG}.$$

From Eq. (1–20),

$$S \leftrightarrow f(G) + c, \qquad (1\text{--}21)$$

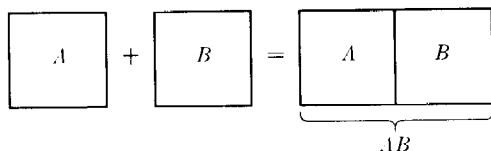where c is an integration constant independent of G and therefore inde-

FIG. 1-3. Systems A and B combined to form AB. All systems are at same temperature.

pendent of the variables on which G depends (e.g., β and V, with N constant). In thermodynamic language, c is independent of the thermodynamic state of a closed system. Experimental information about the entropy always involves a *difference* in entropy between two states (e.g., the entropy change ΔS between T_1 and T_2 at constant N and V), never an absolute value. The constant c in Eq. (1–21) always cancels on taking such a difference. Hence its value is completely arbitrary from an operational point of view. But for convenience and simplicity, we adopt the particular choice $c = 0$ from now on. The connection between this choice and the third law of thermodynamics will be discussed in Section 2–4.

Up to this point we have that $S \leftrightarrow f(G)$, but we do not know the function f. To settle this matter we make use of a thermodynamic property of the entropy, namely its additivity. Specifically, suppose we have two thermodynamic systems A and B at the same temperature and with entropies S_A and S_B. Then if we regard the combined systems (Fig. 1–3) as a new system AB, we have $S_{AB} = S_A + S_B$. This relationship suffices to determine f, as we now show.
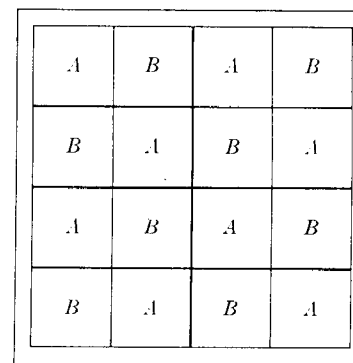
FIG. 1-4. Canonical ensemble of $\mathfrak{N}$ systems, each of type AB.

We first investigate whether the statistical-mechanical quantity G is additive in the above sense. For this purpose we form a canonical ensemble of $\mathfrak{N}$ systems AB (as shown in Fig. 1–4) representative of a thermodynamic (prototype) AB system at temperature T. Heat can flow through all interior walls of the ensemble. The A part of the thermodynamic system is characterized further by N^A and V^A, and the B part by N^B and V^B (A and B are not exponents). In general, the types of molecules may be different in A and B. We have two sets of energy states for the separate systems, $E_1^A, E_2^A, \ldots$ and $E_1^B, E_2^B, \ldots$. If n_j^A stands for the number of A systems in the ensemble in state E_j^A, with a similar meaning for n_j^B, then the number of states of the whole ensemble (Fig. 1–4), or super-

system, consistent with given distributions n^A and n^B is

$$\Omega_t(n^A, n^B) = \frac{(\sum_j n_j^A)!}{\prod_j n_j^A!} \times \frac{(\sum_j n_j^B)!}{\prod_j n_j^B!}, \qquad (1\text{-}22)$$

since the A and B systems are independent of each other (except for energy exchange through the walls). The distributions of interest must satisfy the equations

$$\sum_j n_j^A = \mathfrak{N}, \qquad \sum_j n_j^B = \mathfrak{N},$$

$$\sum_j (n_j^A E_j^A + n_j^B E_j^B) = E_t.$$

The argument from here on is essentially the same as before, so we omit details (Problem 1-3). The three restrictions above require three undetermined multipliers, α_A, α_B, and β, respectively. We note in particular that because of energy exchange between the A and B systems, only one energy equation and one multiplier β are necessary. For the probability that the thermodynamic system AB has its A part in state E_i^A and its B part in state E_j^B, we find

$$P_{ij} = \frac{e^{-\beta E_i^A} e^{-\beta E_j^B}}{Q_A Q_B} = P_i^A P_j^B, \qquad (1\text{-}23)$$

where

$$Q_A = \sum_j e^{-\beta E_j^A}, \qquad Q_B = \sum_j e^{-\beta E_j^B}.$$

This multiplicative property of P_{ij} is of course what we should expect from the form of Eq. (1-22). We deduce from Eq. (1-23) that if two systems are in thermal contact with each other (and therefore have the same temperature), they have the same β. This suggests a close connection between β and T, which we verify below.

For the combined system AB,

$$\begin{aligned}
G_{AB} &= -\sum_{i,j} P_{ij} \ln P_{ij} \\
&= -\sum_{i,j} P_i^A P_j^B (\ln P_i^A + \ln P_j^B) \\
&= -\sum_i P_i^A \ln P_i^A - \sum_j P_j^B \ln P_j^B \\
&= G_A + G_B. \qquad (1\text{-}24)
\end{aligned}$$

That is, G is additive. Also, since $S_{AB} = S_A + S_B$, we have

$$f(G_{AB}) = f(G_A) + f(G_B).$$

Then, from Eq. (1-24),

$$f(G_A + G_B) = f(G_A) + f(G_B).$$

The question before us becomes, then: Given that

$$f(x + y) = f(x) + f(y), \qquad (1\text{-}25)$$

what is the function f? Let us differentiate* Eq. (1-25) with respect to x and y:

$$\frac{df(x + y)}{d(x + y)} \frac{\partial(x + y)}{\partial x} = \frac{df(x + y)}{d(x + y)} = \frac{df(x)}{dx},$$

$$\frac{df(x + y)}{d(x + y)} \frac{\partial(x + y)}{\partial y} = \frac{df(x + y)}{d(x + y)} = \frac{df(y)}{dy}.$$

Hence

$$\frac{df(x)}{dx} = \frac{df(y)}{dy}.$$

This says that a certain function of x is equal to the same function of y. But this is only possible if the function is a constant, say k. Then

$$\frac{df(x)}{dx} = k, \qquad f(x) = kx + a,$$

where a is another constant. But we have to choose $a = 0$ in order to satisfy Eq. (1-25). Therefore, finally, we have found that $f(x) = kx$, and that

$$\begin{aligned}
S \leftrightarrow f(G) &= kG \\
&\leftrightarrow -k \sum_j P_j \ln P_j. \qquad (1\text{-}26)
\end{aligned}$$

Also, from Eq. (1-20),

$$\frac{1}{\beta T} = \varphi(G) = \frac{df(G)}{dG} = k,$$

or

$$\frac{1}{T} \leftrightarrow \beta k, \qquad \frac{1}{kT} \leftrightarrow \beta. \qquad (1\text{-}27)$$

---

* This argument is from Schrödinger, p. 13.

The constant k is still unevaluated at this stage. We have seen that if *any* two systems are in thermal contact, they have the same β and T. Therefore they have the same k. What is more, k is a *universal* constant, since one system of the pair, say A, can be retained and B can be varied over all other possible systems, C, D, E, ... The value of k can thus be obtained once and for all by comparing statistical-mechanical and experimental values of the same property, on any convenient system (A, above). The pressure of an ideal gas is usually used. The numerical value of k depends, of course, on the absolute temperature scale employed. We anticipate from our treatment of an ideal gas in Chapter 4 that $k = +1.38044 \times 10^{-16}$ erg $\cdot$ deg^{-1}, with the conventional kelvin temperature scale. However, the important fact that k is a positive number can easily be checked here in several ways. For example, if we put $\beta = 1/kT$ in Eq. (1–11), differentiate with respect to T, and use the experimental thermodynamic fact that $C_V = (\partial E/\partial T)_{N,V}$ is always positive, we find that k must be positive (Problem 1–4).

We are now in a position to summarize the basic statistical-mechanical equations that can be used to calculate the thermodynamic properties of a closed, isothermal system. In the first place, the probability that the system is in any particular energy state E_j is

$$P_j(N, V, T) = \frac{e^{-E_j(N,V)/kT}}{Q(N, V, T)}, \qquad (1\text{–}28)$$

where

$$Q(N, V, T) = \sum_j e^{-E_j(N,V)/kT}. \qquad (1\text{–}29)$$

We call Q the "canonical ensemble partition function." Because of the association (1–27), the independent thermodynamic variables here turn out to be N, V, and T, which is just the desired set for a closed, isothermal system (see Section 1–3). The entropy is

$$S(N, V, T) = -k \sum_j P_j \ln P_j, \qquad (1\text{–}30)$$

where P_j is given by Eq. (1–28). If we substitute Eq. (1–28) into Eq. (1–30), we find

$$S = \frac{\overline{E}}{T} + k \ln Q = \frac{E}{T} - \frac{A}{T},$$

where the last expression is a thermodynamic one (A is the Helmholtz free energy). Therefore

$$A(N, V, T) = -kT \ln Q(N, V, T). \qquad (1\text{–}31)$$

This equation is particularly useful because A is the "characteristic func-

tion" in thermodynamics for the independent variables N, V, T:

$$dA = -S\, dT - p\, dV + \sum_\alpha \mu_\alpha\, dN_\alpha. \qquad (1\text{–}32)$$

Thus,

$$S = -\left(\frac{\partial A}{\partial T}\right)_{V,N} = kT\left(\frac{\partial \ln Q}{\partial T}\right)_{V,N} + k \ln Q, \qquad (1\text{–}33)$$

$$p = -\left(\frac{\partial A}{\partial V}\right)_{T,N} = kT\left(\frac{\partial \ln Q}{\partial V}\right)_{T,N}, \qquad (1\text{–}34)$$

$$E = -T^2\left(\frac{\partial A/T}{\partial T}\right)_{V,N} = kT^2\left(\frac{\partial \ln Q}{\partial T}\right)_{V,N}. \qquad (1\text{–}35)$$

Hence, if the function $Q(N, V, T)$ is available from Eq. (1–29), differentiation of Q yields S, p, and E. Furthermore, despite the fact that Eq. (1–31) was derived from the study of a closed system, we can make use of the thermodynamic equation (1–32) and $Q(N, V, T)$ to deduce the chemical potential of any component, say i, from

$$\mu_i = \left(\frac{\partial A}{\partial N_i}\right)_{T,V,N_{\alpha \neq i}} = -kT\left(\frac{\partial \ln Q}{\partial N_i}\right)_{T,V,N_{\alpha \neq i}}. \qquad (1\text{–}36)$$

Thus we have a complete set of thermodynamic functions (from which all others can be derived): N, V T; A, E, S, p, μ. Incidentally, whether the averaging bars over E and p in the above equations are dropped or not is optional; it depends on whether one has in mind primarily the thermodynamic or the statistical-mechanical aspect of the equation in question.

The above equations, which allow us to deduce all the thermodynamic properties from Eq. (1–29) for the partition function Q, are general but quite formal. In fact, the reader may feel that these relations are rather useless since, in general, the E_j must be expected to be very difficult to calculate for a system with many molecules. While such an attitude is perhaps justified in complicated cases, there are many systems for which considerable progress of one kind or another can be made. Much of the rest of this book will be devoted to such examples.

For many purposes it is convenient to group together all energy states belonging to the same energy level. Let $\Omega_i(N, V)$ be the number of such states (that is, the degeneracy) for an energy level $E_i(N, V)$. In other words, in the list of energy states E_1, E_2, ..., the same value E_i occurs Ω_i times. Then,

$$Q(N, V, T) = \underset{\text{(states)}}{\sum_j e^{-E_j(N,V)/kT}} = \underset{\text{(levels)}}{\sum_i \Omega_i(N, V)e^{-E_i(N,V)/kT}}. \qquad (1\text{–}37)$$

Also,

$$P \text{ (level)} = \Omega P \text{ (state)} = \frac{\Omega e^{-E/kT}}{Q} \qquad (1\text{-}38)$$

is the probability that the system exists in the energy level E. We have dropped subscripts here to avoid confusion between i and j. Whether a sum such as one of those occurring in Eq. (1-37) is over "states" or "levels" can always be judged by noticing whether or not degeneracies are included as weights for the so-called Boltzmann factors $(e^{-E_j/kT})$.

We have already mentioned that P_j, being proportional to the Boltzmann factor $e^{-E_j/kT}$, falls off exponentially with increasing E_j. We shall discuss essentially this point in more detail in Chapter 3, but in anticipation we should mention here two important extreme cases:

(a) If $T \rightarrow 0$ and the lowest level E_1 is nondegenerate, then

$$Q \rightarrow e^{-E_1/kT}[1 + \Omega_2 e^{-(E_2-E_1)/kT} + \cdots] \rightarrow e^{-E_1/kT}$$

and

$$P_1 \rightarrow 1, \qquad P_j \rightarrow 0, \qquad j = 2, 3, \ldots.$$

That is, in the limit as $T \rightarrow 0$, the system is certain to be found in the lowest energy state. From Eq. (1-30), $S \rightarrow 0$.

(b) If $T \rightarrow \infty$, the relative effect of different E_j's on the Boltzmann factors is washed out, and P_j (state) $\rightarrow$ constant (independent of j); that is, the probability distribution over states becomes uniform. Then $S \rightarrow \infty$, assuming that there is an infinite number of energy states (Problem 1-5).

1-5 Grand canonical ensemble. In this section we suppose that the thermodynamic system of volume V, whose properties we wish to calculate from molecular considerations, is in a large heat bath and is "open" with respect to the molecules in the system. That is, both heat and matter (molecules) can be transported across the walls of the system. The bath provides a reservoir of heat at temperature T and of molecules at chemical potentials $\mu_1, \mu_2, \ldots$. The system is thus characterized by the thermodynamic variables $V, T, \mu_1, \mu_2, \ldots$. The numbers of molecules $N_1, N_2, \ldots$ do not have fixed values, as they do in a closed system, but fluctuate about mean values $\overline{N}_1, \overline{N}_2, \ldots$.

We employ here the same type of argument as for the canonical ensemble: (a) the first postulate permits us to use ensemble averages over mechanical variables in place of time averages on the actual system; (b) by regarding the entire ensemble as an isolated supersystem, we can deduce ensemble-average weighting (probability) factors from the second postulate, in terms of undetermined multipliers; and (c) the significance of the undetermined multipliers, as nonmechanical variables, can then be
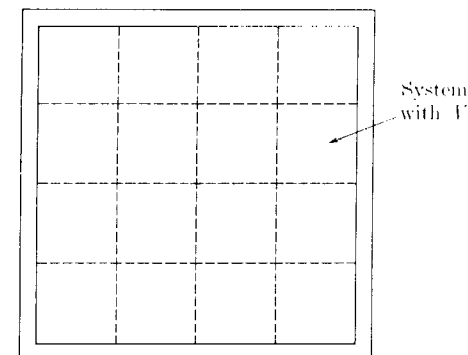
FIG. 1-5. Grand canonical ensemble of $\mathfrak{N}$ systems, each with V.

established by comparing statistical-mechanical and thermodynamic expressions for mechanical variables.

For simplicity, we consider now a one-component system, with V, T, and μ given. As an ensemble (which we call a "grand canonical ensemble") representative of this system, we introduce a lattice (Fig. 1-5) of $\mathfrak{N}$ systems, each with volume V and with walls permeable to molecules (indicated by the dashed lines) and to heat. To establish the desired values of T and μ in each of the $\mathfrak{N}$ systems, we imagine that the whole ensemble is immersed in a giant reservoir at T and μ until equilibrium is reached. We then place walls around the ensemble (solid double lines in Fig. 1-5) that are impermeable to both heat and molecules, and finally remove the ensemble from the reservoir. The ensemble itself is then an isolated supersystem to which the second postulate can be applied. The volume of the supersystem is $\mathfrak{N}V$, and we let E_t and N_t be its total energy and number of molecules.

Since the subsequent details are very similar to those in Sections 1-3 and 1-4, we condense the discussion here. For each value of N, there will be a different set of energy states $E_j(N, V)$. The quantum-mechanical state of the supersystem (ensemble) is specified when we give the value of N and the state $E_j(N, V)$ for each system in the supersystem. In a given state of the supersystem, let $n_j(N)$ be the number of systems which contain N molecules *and* are in the particular energy state $E_j(N, V)$. N can range from zero to infinity (unless there is some upper limit set by the model being used). For a given distribution n, that is, for a set of numbers

$$n_1(0),$$
$$n_1(1), n_2(1), n_3(1), \ldots,$$
$$n_1(2), n_2(2), n_3(2), \ldots,$$
$$\vdots$$

the number of possible quantum states of the supersystem is

$$\Omega_t(n) = \frac{[\sum_{j,N} n_j(N)]!}{\prod_{j,N} n_j(N)!}. \tag{1-39}$$

Acceptable distributions must satisfy the conservation relations

$$\sum_{j,N} n_j(N) = \mathfrak{N}, \tag{1-40}$$

$$\sum_{j,N} n_j(N) E_j(N, V) = E_t, \tag{1-41}$$

$$\sum_{j,N} n_j(N) N = N_t. \tag{1-42}$$

If we let α, β, and γ be the respective undetermined multipliers, the most probable distribution turns out to be (Problem 1-6)

$$n_j^*(N) = \mathfrak{N} e^{-\alpha} e^{-\beta E_j(N, V)} e^{-\gamma N}. \tag{1-43}$$

Again, in principle we can substitute Eq. (1-43) into Eqs. (1-40) through (1-42) and find α, β, and γ as functions of $\mathfrak{N}$, E_t, and N_t. But, instead, we follow a procedure analogous to that used for the canonical ensemble.

From Eqs. (1-40) and (1-43),

$$e^\alpha = \sum_{j,N} e^{-\beta E_j(N, V)} e^{-\gamma N} \tag{1-44}$$

and

$$P_j(N) = \frac{\overline{n_j(N)}}{\mathfrak{N}} = \frac{n_j^*(N)}{\mathfrak{N}} = \frac{e^{-\beta E_j(N, V)} e^{-\gamma N}}{\sum_{i, N'} e^{-\beta E_i(N', V)} e^{-\gamma N'}}, \tag{1-45}$$

where $P_j(N)$ is the probability that a system selected at random from the grand ensemble will contain N molecules and be in the energy state $E_j(N, V)$; or, $P_j(N)$ is the probability that the single prototype thermodynamic system contains exactly N molecules and is in the energy state $E_j(N, V)$. We note that $P_j(N)$ has an exponential dependence on both $E_j(N, V)$ and N. An open system has a definite volume, but both the energy and number of molecules in the system fluctuate. In a closed, isothermal system (canonical ensemble), N is fixed but the energy fluctuates. The magnitude of these fluctuations will be examined in Chapter 2.

From the first postulate, we have the associations

$$E \leftrightarrow \overline{E}(= E_t/\mathfrak{N}) = \sum_{j,N} P_j(N) E_j(N, V), \tag{1-46}$$

$$N \leftrightarrow \overline{N}(= N_t/\mathfrak{N}) = \sum_{j,N} P_j(N) N, \tag{1-47}$$

$$p \leftrightarrow \overline{p} = \sum_{j,N} P_j(N) \left\{ -\left[\frac{\partial E_j(N, V)}{\partial V} \right]_N \right\}. \tag{1-48}$$

These are mechanical variables. To include nonmechanical variables and to evaluate β and γ, we utilize the expression

$$d\overline{E} = \sum_{j,N} E_j(N, V)\, dP_j(N) + \sum_{j,N} P_j(N)\, dE_j(N, V). \tag{1-49}$$

Since we are summing over all values of j and N, $E_j(N, V)$ is in effect a function of V only in the second term on the right. In the first term, we substitute for $E_j(N, V)$ from Eq. (1-45). Then,

$$d\overline{E} = -\frac{1}{\beta} \sum_{j,N} [\gamma N + \ln P_j(N) + \ln \Xi]\, dP_j(N)$$
$$+ \sum_{j,N} P_j(N) \frac{\partial E_j(N, V)}{\partial V}\, dV, \tag{1-50}$$

where

$$\Xi = \sum_{j,N} e^{-\beta E_j(N, V)} e^{-\gamma N}. \tag{1-51}$$

Using [Eq. (1-47)]

$$d\overline{N} = \sum_{j,N} N\, dP_j(N),$$

Eq. (1-50) simplifies to

$$-\frac{1}{\beta} d\left[\sum_{j,N} P_j(N) \ln P_j(N) \right] = d\overline{E} + \overline{p}\, dV + \frac{\gamma}{\beta}\, d\overline{N}. \tag{1-52}$$

We compare this with the thermodynamic equation

$$T\, dS = dE + p\, dV - \mu\, dN, \tag{1-53}$$

and conclude that

$$\mu \leftrightarrow -\frac{\gamma}{\beta}, \tag{1-54}$$

$$T\, dS \leftrightarrow -\frac{1}{\beta} d\left[\sum_{j,N} P_j(N) \ln P_j(N) \right]. \tag{1-55}$$

By the same kind of lengthy argument already employed for the canoni-

cal ensemble, we arrive at the further results

$$S \leftrightarrow -k \sum_{j,N} P_j(N) \ln P_j(N), \qquad (1\text{-}56)$$

$$\frac{1}{kT} \leftrightarrow \beta, \qquad (1\text{-}57)$$

and therefore, from (1-54),

$$\frac{\mu}{kT} \leftrightarrow -\gamma. \qquad (1\text{-}58)$$

The relation (1-56) has the same formal appearance as Eq. (1-30) in the canonical ensemble. In fact, this form for S is quite general (see Problem 1-8, for example).

According to Eq. (1-45), there is only one β (and therefore k) for all values of N. Furthermore, this is the same β as in Section 1-4 for a closed, isothermal system, since a grand ensemble is just an aggregate of canonical ensembles. That is, we can imagine "freezing" the composition of the systems in a grand ensemble by suddenly inserting, between the systems, walls which are heat conducting but impermeable to molecules. Then the original grand canonical ensemble becomes simply a collection of canonical ensembles (in fact, this is the significance of the word "grand") in thermal contact with each other, each characterized by a definite N.

Let us now summarize results for an open, isothermal system whose thermodynamic state is specified by the variables V, T, and μ. The probability that such a system contains N molecules and is in the energy state $E_j(N, V)$ is

$$P_j(N; V, T, \mu) = \frac{e^{-E_j(N,V)/kT} e^{N\mu/kT}}{\Xi(V, T, \mu)}, \qquad (1\text{-}59)$$

where

$$\Xi(V, T, \mu) = \sum_{j,N} e^{-E_j(N,V)/kT} e^{N\mu/kT}. \qquad (1\text{-}60)$$

We call Ξ the "grand partition function." The notation used for P in Eq. (1-59) means that N and j are essentially running indices (the notation P_{Nj} might have been used), while V, T, and μ are independent thermodynamic variables. An alternative form for Ξ is

$$\Xi(V, T, \mu) = \sum_N \left[e^{N\mu/kT} \sum_j e^{-E_j(N,V)/kT} \right]$$
$$= \sum_N Q(N, V, T) e^{N\mu/kT}. \qquad (1\text{-}61)$$

The probability that the system has N molecules, irrespective of the energy state, is

$$P(N; V, T, \mu) = \sum_j P_j(N) = \frac{Q(N, V, T) e^{N\mu/kT}}{\Xi(V, T, \mu)}. \qquad (1\text{-}62)$$

Thus, for example, the average value of N is

$$\overline{N}(V, T, \mu) = \frac{\sum_N N Q(N, V, T) e^{N\mu/kT}}{\Xi(V, T, \mu)}. \qquad (1\text{-}63)$$

If we substitute Eq. (1-59) into Eq. (1-56), we find

$$S = \frac{\overline{E}}{T} - \frac{\overline{N}\mu}{T} + k \ln \Xi = \frac{E}{T} - \frac{N\mu}{T} + \frac{pV}{T},$$

where the last expression is thermodynamic in origin. Hence

$$pV = kT \ln \Xi \ (V, T, \mu). \qquad (1\text{-}64)$$

Now pV is the thermodynamic characteristic function for the variables V, T, and μ:

$$d(pV) = S \, dT + N \, d\mu + p \, dV. \qquad (1\text{-}65)$$

Therefore, from Eq. (1-64), we have the following relations which, together with Eq. (1-64), permit us to calculate all the thermodynamic properties of a system if Ξ is known as a function of V, T, and μ:

$$S = kT \left(\frac{\partial \ln \Xi}{\partial T} \right)_{V,\mu} + k \ln \Xi, \qquad (1\text{-}66)$$

$$N = kT \left(\frac{\partial \ln \Xi}{\partial \mu} \right)_{V,T}, \qquad (1\text{-}67)$$

$$p = kT \left(\frac{\partial \ln \Xi}{\partial V} \right)_{\mu,T} = kT \frac{\ln \Xi}{V}. \qquad (1\text{-}68)$$

The last form of Eq. (1-68) follows from Eq. (1-64) or, on thermodynamic grounds, from the fact that the variables held constant in the derivative are both intensive.

We shall see in Chapter 2 that one can choose an ensemble from which to calculate thermodynamic functions on the basis of convenience, and irrespective of the actual environment of a system (heat bath, constant pressure, etc.). In many problems the grand ensemble is easier to use than the canonical ensemble. When this is the case, the reason is usually either (a) that a mathematically awkward restraint of constant N in the canonical ensemble can be avoided by summing over N (Eq. 1-61), or (b) that

a many-body problem can be reduced to a one-body, two-body, etc., problem by viewing Eq. (1–61) as a power series in the "absolute activity" $\lambda = e^{\mu/kT}$:

$$\Xi(V, T, \mu) = Q(0, V, T) + Q(1, V, T)\lambda + Q(2, V, T)\lambda^2 + \cdots. \quad (1\text{–}69)$$

This is the preferable method in treating an imperfect gas, for example (Chapter 15).

The above discussion is limited to a one-component system, but it can easily be extended to any number of components (Problem 1–7). For example, for two components there will be two equations like (1–42) and two undetermined multipliers, $\gamma_1 = -\mu_1/kT$ and $\gamma_2 = -\mu_2/kT$. Equation (1–59) becomes

$$P_j(N_1, N_2; V, T, \mu_1, \mu_2) = \frac{e^{-E_j(N_1, N_2, V)/kT}\lambda_1^{N_1}\lambda_2^{N_2}}{\Xi(V, T, \mu_1, \mu_2)}, \quad (1\text{–}70)$$

with

$$\Xi = \sum_{N_1, N_2} Q(N_1, N_2, V, T)\lambda_1^{N_1}\lambda_2^{N_2}, \quad (1\text{–}71)$$

where

$$\lambda_1 = e^{\mu_1/kT}, \qquad \lambda_2 = e^{\mu_2/kT}.$$

Also

$$pV = kT \ln \Xi(V, T, \mu_1, \mu_2), \quad (1\text{–}72)$$

$$d(pV) = S\, dT + N_1\, d\mu_1 + N_2\, d\mu_2 + p\, dV, \quad (1\text{–}73)$$

from which we can immediately write the extensions of Eqs. (1–66) through (1–68).

1–6 Microcanonical ensemble. Here we are concerned with an isolated system with given E, V, and N (N again represents a set $N_1, N_2, \ldots$ if the system is multicomponent). The representative ensemble is called a microcanonical ensemble, as stated in Section 1–2. For an isolated system it is difficult to achieve a direct connection between our two postulates and thermodynamics (e.g., we have used variations in $\bar{E}$ for this purpose in Sections 1–4 and 1–5, but here E is constant). The most common procedure for avoiding this difficulty is to introduce, essentially as a new postulate, the equation $S = k \ln \Omega$, where $\Omega(N, V, E)$ is the degeneracy of the energy level E (see Section 1–2). However, a new postulate is not really needed; its introduction is therefore unsatisfactory from a logical point of view. Instead, we derive the properties of a microcanonical ensemble from either the canonical ensemble or the grand ensemble.

First, consider a canonical ensemble. A microcanonical ensemble, as the name is meant to imply, is a degenerate canonical ensemble in which all

systems have (virtually) the same energy. Thus, suppose we start with a canonical ensemble, pick out just those systems with an energy level E, place thermal insulation around each of them, and then remove these systems from the other systems in the canonical ensemble (with energies different from E). As a result of this operation, we have a collection of isolated systems, all with the same N, V, *and* E (a microcanonical ensemble). This degenerate canonical ensemble* may be thought of as being representative of a hypothetical closed, isothermal system that is somehow restrained from having values of E other than $E = \bar{E}$. Another way of saying this is that the only quantum states accessible to the system are those with energy E. In this new ensemble, according to Eq. (1–28), the fraction of systems P_j in a given quantum state (energy E) is proportional to $e^{-E/kT}$. But E is the same for all quantum states, $\Omega(N, V, E)$ in number. Hence P_j is the same for all Ω quantum states. Since $\sum_j P_j = 1$, $P_j = 1/\Omega$. Then, from Eq. (1–30),

$$S(N, V, E) = -k \sum_j P_j \ln P_j = -k\Omega \left(\frac{1}{\Omega} \ln \frac{1}{\Omega}\right)$$

$$= k \ln \Omega(N, V, E). \quad (1\text{–}74)$$

This relation between the thermodynamic S and statistical-mechanical (actually, quantum-mechanical) Ω can then be employed to derive all thermodynamic functions of interest, as we shall see below.

A microcanonical ensemble is also a degenerate grand ensemble: we can pick out of a grand ensemble only those systems with certain prescribed values of N and E. But there is a different, in fact, complementary, way in which Eq. (1–74) can be deduced from a grand ensemble.

In what follows, we have to make use of Section 1–5, which was restricted for simplicity to a one-component system, but the method and result are independent of the number of components. In Section 1–5 we applied the second postulate to the whole grand ensemble, or supersystem (Fig. 1–5). That is, the supersystem itself is an example of an isolated system. The point of view we adopt here is that the supersystem may be regarded, for present purposes, not as an imaginary construct but as a single, very large, *real isolated system*. In this case, the dashed lines in Fig. 1–5, dividing the "supersystem" into "systems," represent *mathematical* rather than physical planes. A "system" (Fig. 1–6) is then an imaginary macroscopic portion, between mathematical planes, of the total or "supersystem." Each "system" is *open* and *isothermal*. Note that

---

* We shall see in Section 2–2 that, because relative fluctuations in E about $\bar{E}$ in a canonical ensemble are virtually negligible in magnitude, this somewhat artificial way of forming a microcanonical ensemble is really unnecessary: canonical and microcanonical ensembles are essentially indistinguishable in any case.
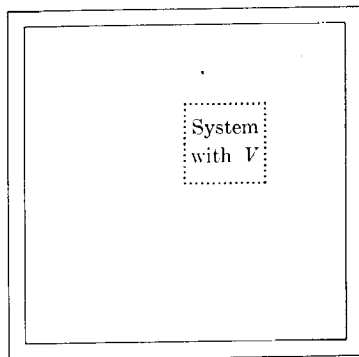
FIG. 1-6. Portion of a larger system forming a smaller, open, isothermal system of volume V.

we cannot use this point of view in connection with a canonical ensemble, because this ensemble requires that each "system" be *closed*. The way we proceed is to employ Eq. (1-39) to find the total number of quantum states Ω_t of the "supersystem" (specified E_t, N_t, $V_t = \mathfrak{N}V$) and relate this number to the entropy S_t of the "supersystem" using (a) the connections with thermodynamics already found for an open, isothermal system (Eqs. 1-59 through 1-68), and (b) the additive property of the entropy, $S_t = \mathfrak{N}S$, where S is the entropy of one "system" in the "supersystem." To evaluate $\ln \Omega_t$, we start with

$$\ln \Omega_t = \ln \sum_n \Omega_t(n) = \ln \Omega_t(n^*) \tag{1-75}$$

and then substitute

$$n_j^*(N) = \frac{\mathfrak{N}e^{-E_j(N,V)/kT}e^{N\mu/kT}}{\Xi}, \tag{1-76}$$

which follows from Eqs. (1-45) and (1-59), into $\ln \Omega_t(n)$ from Eq. (1-39), to give $\ln \Omega_t(n^*)$. We find

$$\ln \Omega_t(n^*) = \mathfrak{N}\ln\mathfrak{N} - \sum_{j,N} n_j^*(N)\ln n_j^*(N)$$

$$= \frac{\mathfrak{N}\overline{E}}{kT} - \frac{\mathfrak{N}\overline{N}\mu}{kT} + \mathfrak{N}\ln\Xi$$

$$= \mathfrak{N}\left(\frac{E}{kT} - \frac{N\mu}{kT} + \frac{pV}{kT}\right) = \frac{\mathfrak{N}S}{k},$$

so that

$$S_t = k \ln \Omega_t, \tag{1-77}$$

in agreement with Eq. (1-74).

Equation (1-74), due to Boltzmann, is possibly the best-known equation in statistical mechanics, mainly for historical reasons. We shall discuss it further in Chapter 2. But let us note here that: (a) it applies to an isolated system; (b) if the ground state of the system is nondegenerate, $\Omega = 1$ and $S = 0$ $(T \to 0)$; and (c) for any isolated system whatever, the more quantum states available to the system, the higher the entropy. This is the origin of qualitative statements which correlate the entropy with "probability," "randomness," "disorder," etc. We shall encounter many examples of such a correlation in the present book. Incidentally, we can appeal to thermodynamics for a quick estimate of the magnitude of Ω to be expected in statistical mechanics in general. That is, since it is found experimentally that* $S = O(Nk)$, $\ln \Omega = O(N)$, and $\Omega = O(e^N) = O(10^{10^{20}})$, an impressively large number.

Let us assume that, from quantum mechanics, we have $\Omega(N, V, E)$ for the system of interest. Our next problem is to calculate all thermodynamic functions, not just the entropy S. But S is the characteristic function for the variables N, V, E:

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \sum_\alpha \frac{\mu_\alpha}{T}dN_\alpha. \tag{1-78}$$

Hence

$$\frac{1}{kT} = \left(\frac{\partial \ln \Omega}{\partial E}\right)_{V,N}, \tag{1-79}$$

$$\frac{p}{kT} = \left(\frac{\partial \ln \Omega}{\partial V}\right)_{E,N}, \tag{1-80}$$

$$-\frac{\mu_i}{kT} = \left(\frac{\partial \ln \Omega}{\partial N_i}\right)_{E,V,N_{\alpha \neq i}} \tag{1-81}$$

The temperature is determined by the dependence of Ω on E. Since we know from thermodynamics that T is positive, we can anticipate that Ω will increase with E for any macroscopic quantum-mechanical system. Clearly, the same statement can also be made about $\Omega(V)$.

In practice, except in very simple systems, $\Omega(N, V, E)$ is not available, and the microcanonical equations (1-74) and (1-79) through (1-81) cannot be utilized. In particular, the restriction to constant energy E is usually a difficulty. This can be avoided by passing to the canonical ensemble.

---

* The notation $O(x)$ means "of order x."

1-7 Other ensembles. Many other ensembles and partition functions are possible and are often useful. For example, for a one-component system, if we start with $\Omega(N, V, E)$ and then sum one at a time over E, N, and V, we obtain* the four relations (we write $1/kT = \beta$ here for convenience),

$$\ln \Omega(N, V, E) = \beta T S = \frac{S}{k}, \tag{1-82}$$

$$\ln \sum_E \Omega(N, V, E)e^{-\beta E} = \text{function of } N, V, \beta = -\beta A, \tag{1-83}$$

$$\ln \sum_N \Omega(N, V, E)e^{\beta \mu N} = \text{function of } V, E, \beta\mu = \beta H, \tag{1-84}$$

$$\ln \sum_V \Omega(N, V, E)e^{-\beta p V} = \text{function of } N, E, \beta p = \beta(TS - pV). \tag{1-85}$$

In Eq. (1-84), H is the "heat content," $E + pV$. The summations over possible values of E and V may be replaced by integrations in most problems, as we shall see. But here we use the present simple notation for convenience. The sum in Eq. (1-83) is the same, except for notation, as the sum over energy levels in Eq. (1-37). Previously we established only Eqs. (1-82) and (1-83), but the other two cases, and those below, can be worked out in detail by the same general methods already used (Problem 1-8).

Continuing, we can also sum two at a time over E, N, and V:

$$\ln \sum_{E,N} \Omega(N, V, E)e^{-\beta E}e^{\beta \mu N} = \text{function of } V, \beta, \mu = \beta p V, \tag{1-86}$$

$$\ln \sum_{E,V} \Omega(N, V, E)e^{-\beta E}e^{-\beta p V} = \text{function of } N, \beta, p$$

$$= -\beta N \mu = -\beta F, \tag{1-87}$$

$$\ln \sum_{N,V} \Omega(N, V, E)e^{\beta \mu N}e^{-\beta p V} = \text{function of } E, \beta\mu, \beta p = \beta E. \tag{1-88}$$

Equation (1-86) is the logarithm of the grand partition function, already encountered. The other two equations are new. Equation (1-87) is particularly important† because it is applicable to a system with the familiar

---

* T. L. HILL, *J. Chem. Phys.* **29**, 1423 (1958). See also A. MÜNSTER (Supplementary Reading list); W. B. BROWN, *Mol. Phys.* **1**, 68 (1958); A. MÜNSTER, *ibid.* **2**, 1 (1959); R. A. SACK, *ibid.* **2**, 8 (1959).

† W. B. BROWN, *Mol. Phys.* **1**, 68 (1958).

set of independent variables N, T, and p. In Eq. (1-87), F is the Gibbs free energy, $A + pV$, equal to $N\mu$ for a one-component system. Finally, we can sum over all of E, N, and V:

$$\ln \sum_{E,N,V} \Omega(N, V, E)e^{-\beta E}e^{\beta \mu N}e^{-\beta p V} = \text{function of } \beta, \mu, p = 0. \tag{1-89}$$

This is an exceptional case, since T, μ, and p appear to be independent variables, whereas we know from thermodynamics that at most two of these variables can be independent. The special treatment necessary for this partition function is provided elsewhere (S. M., Chapters 2 and 3).

The characteristic functions are redefined here, in a more systematic way, as dimensionless quantities. In every case there is an appropriate thermodynamic equation which permits calculation of other thermodynamic functions from knowledge of the partition function. For example, for Eq. (1-83),

$$d(-\beta A) = -E\, d\beta + \beta p\, dV - \beta\mu\, dN, \tag{1-90}$$

or for Eq. (1-87),

$$d(-\beta F) = -E\, d\beta - V\, d(\beta p) - \beta\mu\, dN. \tag{1-91}$$

Equation (1-90) is, of course, just a rearrangement of Eq. (1-32) for a closed, isothermal system.

The reader has perhaps noticed that the characteristic function can be written immediately on inspection of the partition function. The rule is: if we replace Ω by $e^{\beta TS}$, then the characteristic function is the sum of the exponents in the partition function. For example, for Eqs. (1-83) and (1-87), respectively:

$$\beta TS - \beta E = -\beta A,$$
$$\beta TS - \beta E - \beta p V = -\beta F. \tag{1-92}$$

The reason for the existence of this rule will be obvious from our discussion of the thermodynamic equivalence of ensembles in Chapter 2. It depends on the legitimacy of replacing the logarithm of a partition function by the logarithm of its maximum term.

There are two further types of ensembles or partition functions we should mention, since they will be encountered in the applications. First, there are some problems in which the "external variable" V is replaced by another external variable (e.g., length or area) or is supplemented by additional external variables (e.g., volume and area are both external variables). Second, in multicomponent systems, there are cases in which it is helpful or necessary to regard the system as open with respect to some components, but not all.